DOI:
10.1049/iet-com.2016.0675

Document Version
Peer reviewed version

Citation for published version (APA):
Naslcheraghi, M., Ghorashi, S. A., & Shikh-Bahaei, M. (2017). Full-Duplex Device-to-Device Communication for Wireless Video Distribution. *IET Communications*, *11*(7), 1074-1081. https://doi.org/10.1049/iet-com.2016.0675

Full-Duplex Device-to-Device Communication for Wireless Video Distribution

Mansour Naslcheraghi[1], Seyed Ali Ghorashi[1,2,*], Mohammad Shikh-Bahaei[3]

[1]Cognitive Telecommunication Research Group, Department of Electrical Eng., Shahid Beheshti University G. C., Tehran, Iran
[2]Cyberspace Research Institute, Shahid Beheshti University G. C., Tehran, Iran
[3]Centre for Telecommunications Research, King's College London, UK
Email: m.naslcheraghi@mail.sbu.ac.ir, a_ghorashi@sbu.ac.ir, m.sbahaei@kcl.ac.uk

Abstract: Spectrum scarcity and dramatically increasing demand for high data rate and high quality video live streaming are of future cellular network design challenges. As a solution to this problem, cache-enabled cellular network architecture has been recently proposed. Device-to-Device (D2D) communications can be exploited for distributed video content delivery, and devices can be used for caching of the video files. This can increase the capacity and reduce the end-to-end delay in cellular networks. In this paper, we propose a new scheme for video distribution over cellular networks by exploiting full-duplex (FD) radios for D2D devices in two scenarios; i) two nodes exchange their desired video files simultaneously, and ii) each node can concurrently transmit to and receive from two different nodes. In the latter case, an intermediate transceiver can serve one or multiple users' file request(s) whilst capturing its desired file from another device in the vicinity. Mathematical expressions along with extensive simulations are used to compare our proposed scheme with a half-duplex (HD) scheme to show the achievable gains in terms of sum throughput, active links, and delay. We will also look into the energy cost for achieving the improvements provided by operation in FD mode.

1. Introduction

Mobile video traffic accounted for 55% of total mobile traffic in 2015, because video content on smart mobile devices needs higher data rates than any other mobile data types. According to Cisco's annual report, mobile video traffic is expected to generate three-quarters of the whole mobile data traffic by 2020 [1]. Hence, increasing demand for high quality video in cellular networks on one hand, and spectrum scarcity on another hand, have spurred researchers' attention to cache-enabled cellular network architectures [2]. The key idea in these systems is using helper nodes instead of cellular infrastructures to deliver the desired content to user devices. Helper nodes are generally categorised to small base stations (SBSs) [3] and user mobile devices which are used for D2D communications. In such systems, a user's mobile device with considerable available storage, caches a number of popular video files based on a particular caching policy utilised and controlled by cellular network base station (BS). Then, it can serve other users who request the cached files via D2D communications.

Video cellular caching architectures in cooperation with D2D communications have been investigated in a variety of scenarios in recent years. Information theoretic bounds for single-hop D2D caching networks are obtained in [4] under arbitrary demand for certain deterministic and random caching policies. In such systems, contents can be placed on collaborative nodes (user devices) formerly, either according to a predefined policy (reactive caching) [5], or more intelligently, according to statistics of the user devices' interests (proactive caching) [6]. Then, users' demand can be served via D2D communications. In contrast with device-centric caching policy, cluster-centric caching policy is proposed in [7] to maximise area spectral efficiency. The main idea in [7] is to place the content in each cluster in order to maximise the collective performance of all devices. Content broadcasting approaches are proposed in [8] to offload data via D2D collaboration, by focusing on the joint optimisation of content transmission rate and device's relay duration [8]. Since candidate users for D2D collaboration may avoid participating, in order to encourage users to collaborate in D2D communications and share their cached contents, an incentivizing mechanism is proposed in [9]. Distributed infrastructure-assisted data offloading algorithms along with proactive caching are also investigated in [10] to increase system performance in terms of average cache capacity and waiting time. Cross-layer resource allocation methods are also investigated for supporting video over wireless in multiuser scenarios [11]. It is shown that quality-aware resource allocation can improve video services in wireless networks.

Thus far, the existing literature related to cellular content caching and delivery networks utilises half-duplex (HD) communication for both caching and delivery processes. However, recent advances in full-duplex (FD) radio design [12], materialised by advanced signal processing techniques that can suppress self-interference (SI) at the receiver, have enabled simultaneous transmission and reception over the same frequency band. From theoretical point of view, FD communication can potentially double the spectral efficiency of a point-to-point link, providing SI is entirely cancelled [13]. Exploiting full-duplex communication in a multi-node network can provide even higher throughput gains if multi-node interference is controlled through full-duplex MAC and network layer protocols. FD-D2D not only increases the spectral efficiency, but also supports the required levels of the quality of service (QoS) for 5G wireless networks [14]. A spectrum sharing mechanism in [15] is proposed to support QoS of cellular users with coexistence of the D2D pairs. Full duplex communication can also reduce the end-to-end delay by allowing parallel transmission and reception/sensing by each node. Delay reduction can be specifically achieved through enabling relay nodes to receive data or video files from one device/node and simultaneously transmit data/video files to another device/node. Improving MAC layer delay for video applications has also been analysed in the literature [16]. Regarding relay capability of nodes, HD/FD relaying schemes based on the channel state information (CSI) at both transmitter and receiver with incorporating power is investigated in [17] to improve system performance.

In this paper, we exploit FD-D2D communications capability for cellular video distribution systems by enabling FD radios for user devices. In cellular caching systems [2,5,6], content delivery via D2D communications is based on HD communication in which each device can either receive its desired content or can serve for another user's demand. However, in this work, we let user devices: a) operate in bidirectional FD mode in which two users exchange data simultaneously, and b) concurrently transmit to and receive from two different nodes at the same frequency. We use mathematical analysis and extensive simulations to evaluate the improvements provided by

FD-capable D2D communications on the system throughput, the number of activated links, and end-to-end delay. Also the tradeoff between energy-efficiency (EE) versus spectral-efficiency (SE) is illustrated for two scenarios of HD-D2D and FD-D2D communications. Simulation results show that even under high interference regime, the FD-enabled cellular caching system outperforms the traditional HD-D2D cashing systems in terms of delay and total average system sum throughput. To the best of our knowledge, this is the first time that FD-D2D communication and collaboration is exploited for video content delivery through distributed caching in cellular systems.

The rest of paper is structured as follows. In Section 2 system model as well as a potential D2D communication graph are introduced. In section 3, throughput analysis for the proposed FD-enabled cellular system is provided. In section 4 simulation methodology and results are explained and conclusions and future works are presented in section 5.

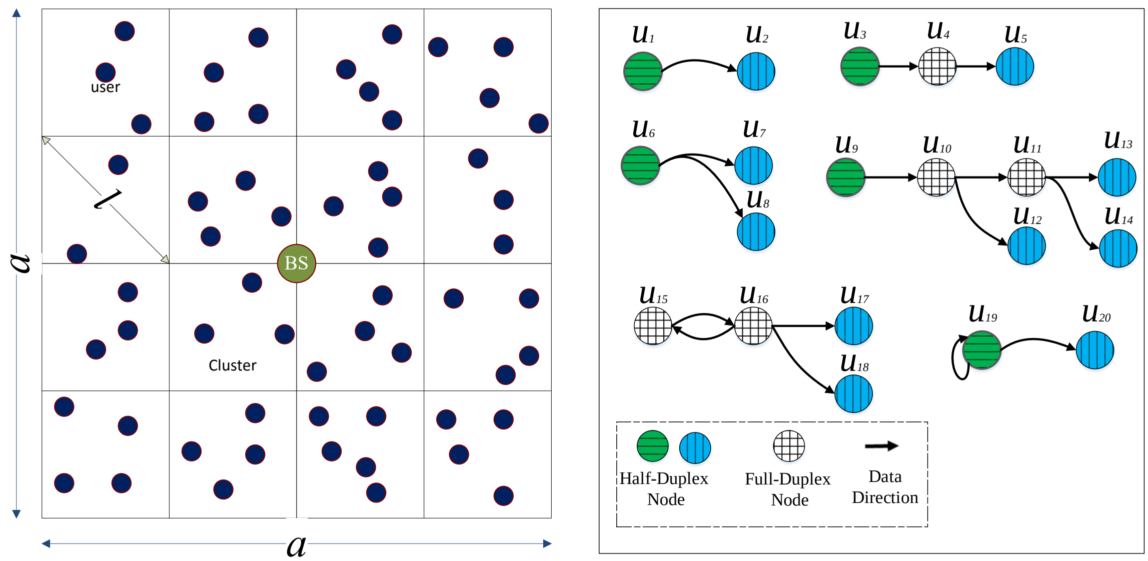
2. SYSTEM MODEL

We consider a cellular network with a single square cell of size a (Fig. 1(a)), one BS and n users that are uniformly distributed within the cell area. Providing that inter-cell interference is negligible or cancelled out, the analysis can be applied to multi-cell scenarios. We divide the whole cell area into virtual equal sized square clusters and neglect co-channel interference and neighbouring cell users' influence for the sake of simplicity. We also assume that SI cancellation allows the FD radios to transmit and receive simultaneously over the same frequency band.

We assume D2D and cellular communications take place in a different frequencies and therefore there is no interference between D2D pairs and cellular users. However, since all D2D pairs in all clusters share the same resource blocks, inter- and intra-cluster interferences will be taken into account. To control and mitigate the mutual interference between the D2D pairs, efficient power control and resource allocation schemes can be applied [19]. Denote the set of popular video files as $M = \{f_1, f_2, ...f_m\}$. We use Zipf distribution for modelling the popularity of video files [18]. Each file has a popularity based on the Zipf distribution and we assume that files popularity and Zipf distribution behaviour remain nearly constant during caching process. Users do not behave proactively, which implies that they would not cache video files in response to the request by users in the cluster, i.e., contents are cached in advance according to a caching policy. It is also assumed that each user has sufficient storage to cache a file from the library. Under these assumptions and for notational simplicity, a caching policy is applied among the users in which each user caches exactly one file from the library and there is no overlap in caches. According to this policy, we assign k most popular files to k distinct users inside the cluster, i.e., user u_i caches the ith file from the library. Thus, the strategy of file assignments and searching the caches for users' desired files within a cluster, is fully controlled by the BS. Each user randomly requests a file from the library, according to Zipf distribution. A pair of users/devices (u_i, u_j) can potentially initiate a D2D communication for video file transfer, providing that the distance between u_i and u_j is less than a threshold l (Fig. 1(a)) and one of them finds its desired file in the other's cache.

From observations and analysis of the frequency of video file statistics in a long period of time [18], the approximate formula for the popularity of file i can be defined by:

$$f_i = \frac{\frac{1}{{\gamma_r}^i}}{\sum\limits_{j=1}^{M} \frac{1}{{\gamma_r}^j}}, \tag{1}$$



(a) (b)

Fig. 1. *System model and random D2D communications graphs, (a) Single cell with equal sized square clusters, (b) Samples of a D2D communications graph.*

where M indicates the library size and γ_r is the Zipf exponent which is an indicator of popularity of the files. In this paper we refer to condition $\gamma_r \geq 1$ as high content reuse regime (high redundancy in files requests, i.e., few files account for the majority of requests), and $\gamma_r \leq 1$, as low content reuse regime (low redundancy in files requests, i.e., popularity of files converges to uniform distribution).

We expect that FD-enabled D2D nodes increase the number of active links inside a cluster and thereby increase the number of satisfied users, whilst they decrease the delay in accessing video files. Given a FD-enabled D2D transmitter, multiple links concurrently can be activated. Fig. 1(b) illustrates the layout of a D2D communications graph that may occur inside a cluster. This graph is constructed based on the content cached in users' devices. We define a directed edge from u_i pointing to u_j if u_j requests a file that has been cached previously by u_i, e.g. u_1 holds a file requested by u_2 and u_3 holds a file requested by u_4. Since we have assumed that each user can make only one request, there will be at most one incoming link to the user and one or multiple outgoing links from the user and no data is relayed over multiple hops. There are two different possible configurations for both HD and FD communications. For HD communications, we have: a) a unicast D2D transmission in which one user targets one receiver, e.g. u_1 targets u_2, and b) a multicast D2D transmission where multiple users can benefit from one transmission inside a cluster, e.g. u_6 targets both u_7 and u_8 in Fig. 1(b). We assume that all users are scheduled to multicast signals. For FD communications, we have: a) a bidirectional FD transmission mode in which two users exchange files simultaneously, e.g. u_{15} and u_{16}, and b) FD radios transmit to and receive from two different nodes, e.g. u_4, u_{10}, and u_{11}. Also, there is a probability that a user, that finds its desired file in its own cache (we call this user as a self-request), can potentially make a D2D communication to serve for other user's demand, e.g. u_{19}.

3. ANALYSIS

We assume that BS can obtain the CSI of all potentially collaborative D2D links. To achieve this, required resources are assigned to potential D2D pairs by the BS. Therefore, BS can calculate the rate of D2D links after estimating the respective signal-to-interference ratio (SIR) values. Subsequently, it can schedule the links to optimise the system performance. If d ($d \leq l$) denotes the Euclidian distance between transmitter u located in u_0 and receiver v located in v_0 inside the cluster, i.e.,

$$d = \|u_0 - v_0\|, \tag{2}$$

the received power, P_v, at receiver v from transmitter u can be written as:

$$P_v = P_t.H_{uv}.d^{-\alpha}, \tag{3}$$

where P_t is the transmitted power of u. The channel is modelled as Rayleigh fading, and thus H_{uv} is an exponentially distributed random variable which represents the Rayleigh fading channel power gain with mean one, i.e., $H_{uv} \sim \exp(1)$, and α is the path loss exponent. We define Ω_T^c as the set of potential D2D transmitters inside the cluster c. Hence, the set of all potential D2D transmitters in the cell, is defined as:

$$\Phi = \bigcup_{c \in \Psi} \Omega_T^c, \tag{4}$$

where Ψ denotes the set of all clusters in the cell. The receiver v may suffer from both intra- and inter-cluster interferences due to the potentially simultaneous transmissions in all clusters. Intra-cluster interference at user v can be defined as

$$I_{intra-cluster} = \sum_{z \in \Omega_T^c \backslash u_0} P_t.H_{zv}.\|z - v_0\|^{-\alpha}, \tag{5}$$

where z denotes the location of a potential transmitter inside the cluster. And inter-cluster interference can be defined as

$$I_{inter-cluster} = \sum_{z \in \Phi \backslash \Omega_T^c} P_t.H_{zv}.\|z - v_0\|^{-\alpha}, \tag{6}$$

where backslash in the summation index of eqs. (5) and (6) is used to indicate that transmitter is excluded from the D2D link u-v and set of all transmitters inside the cluster, respectively. Finally, the measured average SIR at receiver v can be written as

$$SIR = \frac{P_v}{I_{intra-cluster} + I_{inter-cluster}}. \tag{7}$$

We assume that the system operates in the interference limited regime (i.e., the background noise is negligible compared to the interference) except in subsection 3.3, where we will present tradeoff analysis in the system.

3.1. System Sum Throughput

The rate of link i with SIR_i can be calculated as:

$$R_i = \log(1 + SIR_i) \quad \text{bits/s}. \tag{8}$$

FD-D2D video content delivery is used to maximise the expected throughput by activating more links between users. However, any potential throughput gain over FD-D2D content delivery depends on the efficiency of the mechanisms deployed for controlling intra- and inter-cluster interferences. To have a fair comparison between the two transmission modes (HD and FD), we define transmission efficiency coefficient η_δ, as the total average rate for a scheduled D2D transmission:

$$\eta_\delta = \frac{\bar{R}_\delta}{N_\delta}, \tag{9}$$

where $\delta \in \{FD, HD\}$ is the operation mode. N_δ is the number of scheduled D2D transmissions, and $\bar{R}_\delta$ is the maximum achievable (average) sum rate for N_δ D2D transmissions and can be written as: $\bar{R}_\delta = \sum_{i \in L} R_i$, where L denotes the set of all active links inside the cluster. For a cluster index c, given a limit on intra- and inter-cluster interferences imposed on a D2D link, $N_{\delta,c}$ number of simultaneously active D2D transmitters per cluster are scheduled. Hence, the problem can be formulated by

$$\begin{aligned} &\underset{l}{maximise} \sum_{c \in \Psi} \eta_{\delta,c} \\ &s.t. \\ &N_{\delta,c} \leq N_{th}, \end{aligned} \tag{10}$$

where $\eta_{\delta,c}$ is the total average rate per successful D2D transmission of cluster c, and $c \in \Psi$ indicates that we are taking summation over all clusters within the cell. The constraint on the maximum number of D2D transmitters (N_{th}) inside the cluster is set in such a way that both FD and HD communications can take place within a cluster, whilst limiting inter-node (intra- and inter-cluster) interferences. We will describe this constraint further in the next section, where we solve the optimisation problem in (10) through simulations.

3.2. Delay

Now, we calculate the experienced total average delay in the system by taking into account both cellular user equipments (CUEs) and D2D users (DUEs). We first consider self-request users that obtain their desired files with zero delay. We define X as a Zipf-distributed random variable. Denoting $P(X_i = i)$ as probability that user i demands for the ith file from the library, the probability that k users inside the cluster can find their desired files on their own cache is:

$$P_{self}(k) = \sum_{i=1}^{k} P(X_i = i), \tag{11}$$

from (1) it can be written as:

$$P_{Self}(k) = \sum_{i=1}^{k} f_i, \tag{12}$$

where f_k is the popularity of file k. As the distribution of users' positions within the cell area follows uniform distribution, the probability $P_{self}(k)$ in eq. (11) or (12) is the same for all clusters. However, $P_{self}(k)$ depends on the number of users that are randomly distributed within a cluster. Hence, for a given cluster with k randomly selected users, the expected number of self-request users can be calculated as:

$$E[S] = \sum_{k=0}^{n} \sum_{i=1}^{k} f_i . \Pr(K = k), \tag{13}$$

where $Pr(K = k)$ is the probability that there are k users inside the cluster. Since the distribution of users' locations is assumed to be uniform, and the whole cell area is divided into equally-sized virtual square clusters (Fig. 1(a)), the number of users inside a cluster (k) would be a binomial random variable with parameters n and $l^2/2a^2$, i.e., $K = B(n, l^2/2a^2)$:

$$\Pr(K = k) = \left(\begin{array}{c} n \\ k \end{array} \right) \left(\frac{l^2}{2a^2} \right)^k \left(1 - \frac{l^2}{2a^2} \right)^{n-k}. \tag{14}$$

Hence, the total number of self-request users can be calculated as:

$$n_{self} = \frac{2a^2}{l^2} E[S], \tag{15}$$

where $2a^2/l^2$ is the number of clusters. Now we consider DUEs that achieve their desired files via D2D communication in either HD-D2D or FD-D2D mode. We denote the times needed to download a file via traditional cellular communication and D2D collaboration by T_{BS} and T_{D2D}, respectively. Denoting τ as the total download time for all potential D2D pairs and CUEs within the cell, we can write:

$$\tau = (n - n_{self} - L_\delta) T_{BS} + L_\delta T_{D2D}, \tag{16}$$

where n denotes total number of uniformly distributed users within the cell area, and L_δ is the number of users that can achieve their desired files via D2D communications in either HD or FD collaborations. Hence, the problem can be formulated as:

$$\begin{array}{ll} \underset{l}{minimise} & \tau \\ s.t. & \\ \quad s_1 \leq S \leq s_2, & \\ \quad R_{D2D} \leq R_0, & \\ \quad R_{CUE} \leq R_1, & \end{array} \tag{17}$$

where s_1 and s_2 denote the minimum and maximum file sizes in the library, respectively. R_0 and R_1 indicate the maximum data rates in D2D and cellular links that are defined by the cellular network infrastructure.

3.3. Energy-Efficiency and Spectral-Efficiency Tradeoff

Now we investigate the tradeoff between energy-efficiency (EE) and spectral-efficiency (SE) by focusing on a typical random cluster with HD-D2D and FD-D2D communications inside it. In Fig. 2, ζ is the cluster size, d_{ij}, d_{jk} and d_{ik} are the respective distances between D2D users u_i, u_j and u_k. h_{ij}, h_{jk} and h_{ik}, respectively, denote Rayleigh fading channel gains (as was described in section 2). p_{ti} corresponds to the transmit power of user u_i. We focus on the D2D communications graph inside this cluster which is a typical scenario in D2D video content delivery networks as described in section 2. Video content delivery in HD-D2D communication mode is carried out in two consecutive time slots, namely, $T1$ and $T2$, where the length of time slots are assumed to be the same. And communication in FD-D2D mode is performed over entire time frame T, where $T = T1 + T2$. In HD-D2D communication, u_i transmits the desired content of user u_j in the first time slot $T1$ and then u_j transmits the desired content of user u_k in the second time slot $T2$. In FD-D2D communication, user u_j receives its desired content from user u_i and concurrently transmits the desired content of user u_k over the entire time frame T.

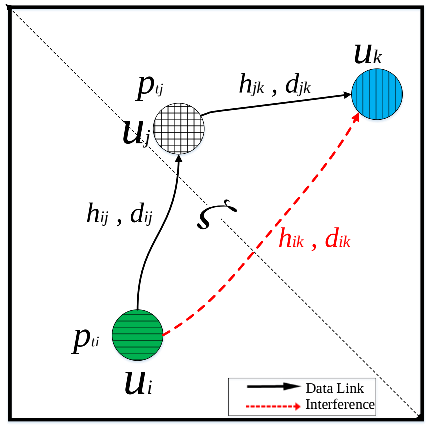


Fig. 2. *A typical random cluster with D2D communications.*

As shown in the Fig. 2, there is no interference imposed on user u_i and u_j. Therefore for analyzing this specific example, we will consider noise as well as interference, in contrast with the interference limited assumption of the analysis in the previous sections. However, for the user u_k there are two possible scenarios; a) user u_j operates in FD-D2D mode, and due to simultaneous transmissions of users u_i and u_j, there will be an interference on user u_k due to transmission by user u_i. b) user u_j operates in HD-D2D mode, and there is no interference on user u_k. Now we define signal-to-interference-plus noise ratio (SINR) for the users u_j and u_k by considering user u_j in FD and HD modes, respectively, as follows

$$\gamma_{ij} = \frac{p_{ti} h_{ij} d_{ij}^{-\alpha}}{\sigma^2}, \tag{18}$$

$$\gamma_{jk}^{FD} = \frac{p_{tj} h_{jk} d_{jk}^{-\alpha}}{p_{ti} h_{ik} d_{ik}^{-\alpha} + \sigma^2}, \tag{19}$$

$$\gamma_{jk}^{HD} = \frac{p_{tj} h_{jk} d_{jk}^{-\alpha}}{\sigma^2}, \tag{20}$$

where α is the path loss exponent as in eq. (3) and σ^2 is additive white Gaussian noise (AWGN) power with zero mean. We formulate the tradeoff based on the analysis in recent works [**?**, **?**]. The overall SE for both HD-D2D and FD-D2D systems can be respectively defined as

$$\Upsilon_{SE}^{HD}(p_{ti}, p_{tj}) = \frac{1}{2}\log_2(1 + \gamma_{ij}) + \frac{1}{2}\log_2(1 + \gamma_{jk}^{HD}), \tag{21}$$

$$\Upsilon_{SE}^{FD}(p_{ti}, p_{tj}) = \log_2(1 + \gamma_{ij}) + \log_2(1 + \gamma_{jk}^{FD}). \tag{22}$$

Now, the system energy efficiency for HD-D2D and FD-D2D can be, respectively, defined as

$$\Upsilon_{EE}^{HD}(p_{ti}, p_{tj}) = \frac{\frac{1}{2}\log_2(1 + \gamma_{ij})}{p_{T1}(p_{ti})} + \frac{\frac{1}{2}\log_2(1 + \gamma_{jk}^{HD})}{p_{T2}(p_{tj})} \quad [bits/Joule], \tag{23}$$

$$\Upsilon_{EE}^{FD}(p_{ti}, p_{tj}) = \frac{\Upsilon_{SE}^{FD}(p_{ti}, p_{tj})}{P_T(p_{ti}, p_{tj})} \quad [bits/Joule], \tag{24}$$

where, $P_{T1}(p_{ti})$ and $P_{T2}(p_{tj})$ respectively denote the consumed powers in the first and second time slots, in HD-D2D mode. $P_T(p_{ti}, p_{tj})$ is the total consumed power over the entire time frame T in FD-D2D mode. These can be calculated as $P_{T1}(p_{ti}) = \frac{1}{2}p_{ti} + 2p_c$, $P_{T1}(p_{tj}) = \frac{1}{2}p_{tj} + 2p_c$, and $P_T(p_{ti}, p_{tj}) = p_{ti} + p_{tj} + 3p_c$, and p_c denotes the power consumption in the circuitry and is regarded as constant. Our objective is to find the optimal p_{ti} and p_{tj} to maximise Υ^{δ}_{EE} under the constraints of SINR and transmission powers. δ is representative of the operation mode as defined in subsection 3.1. Since the SE is the function of channel link SINRs, we shall implicitly consider spectral efficiency as a constraint in the following optimisation equation. With this assumption, we shall maximise energy efficiency, hence, the corresponding optimisation problem for HD-D2D and FD-D2D systems can be formulated as

$$\begin{array}{ll}\max\limits_{p_{ti},p_{tj}} & \Upsilon^{\delta}_{EE} \\ s.t. & \\ \quad \gamma_{th} \leq \gamma_{ij}, \gamma^{\delta}_{jk}, & \\ \quad 0 < p_{ti} \leq \widehat{p}_{ti}, & \\ \quad 0 < p_{tj} \leq \widehat{p}_{tj}, & \\ \quad d_{ij}, d_{jk}, d_{ik} \leq \zeta, & \\ \quad \zeta \leq \widehat{\zeta}, & \end{array}, \tag{25}$$

where γ_{th} is the minimum required SINR for successful demodulation and decoding, $\widehat{p}_{ti}$ and $\widehat{p}_{tj}$ are the maximum allowed transmission powers of nodes u_i and u_j, respectively. In the optimisation problem in eq. (25), Υ^{δ}_{EE} is non-convex with respect to p_{ti} and p_{tj} and analytical solution of the problem is complex. In this paper, we solve this problem numerically through simulations. In addition to the achieved gain in spectral efficiency, we will also improve the video content delivery performance by reducing delay in download time.

3.4. Implementation Challenges

There are some challenges in the implementation of D2D cache enabled networks; one is that some user devices have basic physical capabilities. For such cases, low bandwidth visual data transmission using methods such as compressed sensing and dynamic visual sensing may be incorporated. Another implementation challenge is related to varying video content popularity. Advanced cognitive learning methods can be exploited to update the cache content in a timely fashion. Moreover, the users need to be incentivized to collaborate in such a D2D content delivery scheme, since they have to use their vital resources in this collaboration.

4. SIMULATION RESULTS

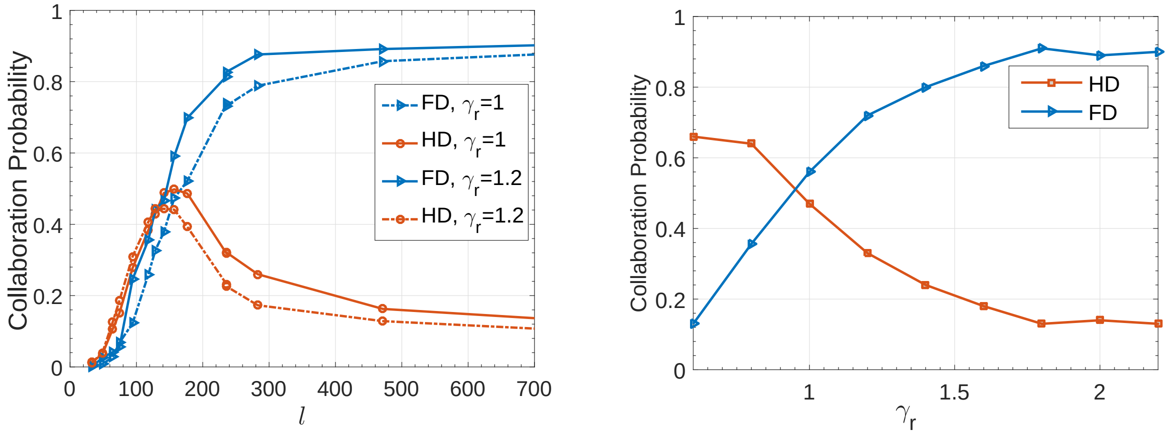
In this section, we provide Monte-Carlo simulations to evaluate the performance of the proposed FD caching system. In our simulations, all users remain stationary and mobility is not taken into account. The rest of simulation parameters are shown in the Table 1. In each iteration of the simulation, a predefined file placement is established such that the k most popular files are assigned to k users inside the cluster, then each user makes a request randomly according to Zipf distribution. According to the content of caches assigned in advance by the BS, all potential D2D communications are identified. We have assumed that all requests taken place independently at the same time. Assuming that the BS is aware of all CSIs between D2D pairs, we first schedule all potential

D2D links in all clusters to obtain the corresponding SIR values. Then, measured SIRs are used to calculate the achievable rate of each potentially active D2D transmitter in both FD and HD modes. As mentioned in Table 1, we choose two D2D transmitters (i.e., $N_{th} = 2$) that can generate the highest rate among all potential D2D transmitters inside a cluster. Setting parameter N_{th} to one, would imply allowing only one active transmitter per cluster, in which case there will be no FD communication inside the cluster. Therefore, we set this value to be larger than one which allows us to compare FD-capable system with its HD counterpart. Optimal setting of the upper-limit (N_{th}) in eq. (10), requires full analysis of sum throughput versus intra-cluster interference when multiple users are active, which is beyond the scope of this paper. For the sake of simplicity and a fair comparison between FD and HD scenarios whilst achieving minimum intra-cell interference, we enable maximum of two active transmitters per cluster, by setting $N_{th} = 2$.

Table 1 Simulation Parameters

Parameter	Values	Description
n	500	Number of users within cell area
M	1000	Size of library
γ_r	[1 2.2]	Zipf exponent
σ^2	-174 dbm/Hz	Background noise
α	2.6	Path loss exponent
P_t	23 dBm	User maximum transmit power
$(s1, s2)$	(10 35) MB	Size of video files constraints in eq. (17)
$\widehat{p}_{ti}, \widehat{p}_{tj}$	23 dBm	Maximum transmit power of the users in eq. (25)
p_c	2 dB	Circuitry consumed power
γ_{th}	25 dB	SINR threshold
$\widehat{\zeta}$	100 m	Maximum collaboration distance constraint in eq. (25)
h_{ij}, h_{jk}, h_{ik}	exp(1)	Rayleigh fading channel, i.e., $h \sim \exp(1)$
N_{th}	2	Number of established nodes per each cluster
R_0	50 Mbps	D2D link rate (eq.
R_1	150 kbps	Cellular link rate
T_{D2D}	4.18 sec	Time needed to download file via D2D link
T_{BS}	1635 sec	Time needed to download file via cellular link
Shadow fading	4 dB	Log-normal Shadow fading with standard deviation of 4 dB

Now, we show that the majority of expected nodes in the proposed system are in collaboration with FD nodes. The collaboration probabilities of FD and HD nodes can be calculated as $P_{FD} = \frac{N_{FD}}{N_{pot}}$ and $P_{HD} = \frac{N_{HD}}{N_{pot}}$, respectively, where N_{FD} and N_{HD} are the expected numbers of FD and HD transmitters, respectively. N_{pot} is the expected number of all potential D2D collaborating transmitters inside the cluster. To obtain collaboration probabilities, we perform simulations for a typical cluster, which is a randomly chosen cluster within the network. Fig. 3 demonstrates the impact of cluster size l and Zipf exponent γ_r on collaboration probabilities. Fig. 3(a) shows the probability that a potential D2D transmitter inside a cluster is attributed to an FD or an HD node. As can be seen in Fig. 3(a), for lower values of l, the probabilities corresponding to HD and FD links are almost the same. This is due to low-density clusters for smaller l. By increasing D2D collaboration threshold l, the number of users inside a cluster increases and consequently the number of users that collaborate with FD nodes, increases, and the number of users that collaborate with HD nodes, decreases. Fig. 3(b) demonstrates the impact of Zipf exponent γ_r on the aforemen-



(a) (b)

Fig. 3. *FD/HD collaboration probability, (a) Impact of D2D collaboration threshold l, (b) Impact of Zipf exponent γ_r.*

tioned probabilities. As can be seen from Fig. 3(b), the higher the value of γ_r (high redundancy in user requests, the higher is the FD collaboration probability. The important fact in Fig. 3(b) is that the probability of FD collaboration is greater than HD collaboration for $\gamma_r \geq 0.95$. Hence, we investigate the effectiveness of FD capability on other system performance factors, i.e. average rate and total download time, by assuming high redundancy in file requests. For $\gamma_r \leq 0.95$, incentive mechanisms can be used to increase FD collaboration probability.

The expected number of links in FD-enabled and HD systems are shown in Fig. 4. To obtain this figure, we have taken into account all potential D2D links within the cell that are connected to the potential FD or HD nodes. Similar explanations to those in Fig. 3 are valid for Fig. 4.

Now, we implement the optimisation problem in eq. (10) to maximise the system sum throughput by setting $N_{th} = 2$ and varying the cluster size l. Fig. 5 shows the total average rate of the FD-only D2D network, by taking into account inter- and intra-cluster interferences. Optimal value of l corresponds to the maximum average sum-rate of the system. This can be interpreted as more D2D communications and thereby higher frequency reuse. Although the number of virtual clusters increases at lower ranges of l (we expect that the frequency reuse increases as well), nevertheless, the probability that clusters are of low density or no D2D candidates found therein, increases too. This can be interpreted as the fact that probability of finding a user's desired file inside the cluster decreases when the node density decreases. As the number of virtual clusters in the cell decreases, the frequency reuse decreases, too. As shown in Fig. 5, incorporating FD-enabled nodes can improve the average gain in sum throughput at optimal l. Alongside the considerable improvements in system throughput at or near the optimal l with a specific Zipf exponent, improvements of the frequency reuse in FD-enabled system is more underlined for higher values of l.

To implement the optimisation problem in eq. (17), we consider all CUEs and DUEs within the cell area. According to the size of files and link capacity as shown in Table 1, the average times needed to download a file via a D2D link and cellular link are approximately 4.18 and 1634 seconds, respectively, i.e., $T_{D2D} = 4.18$ and $T_{BS} = 1635$ in eq. (16). The results of the

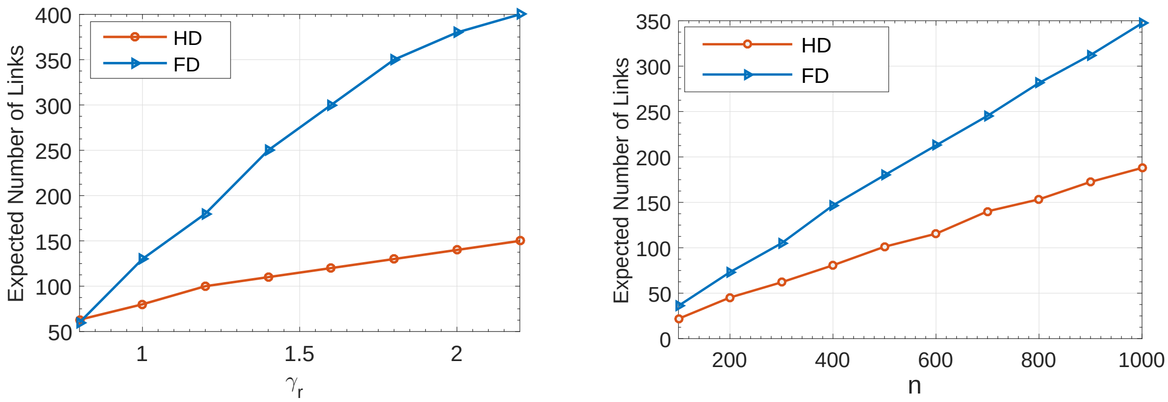


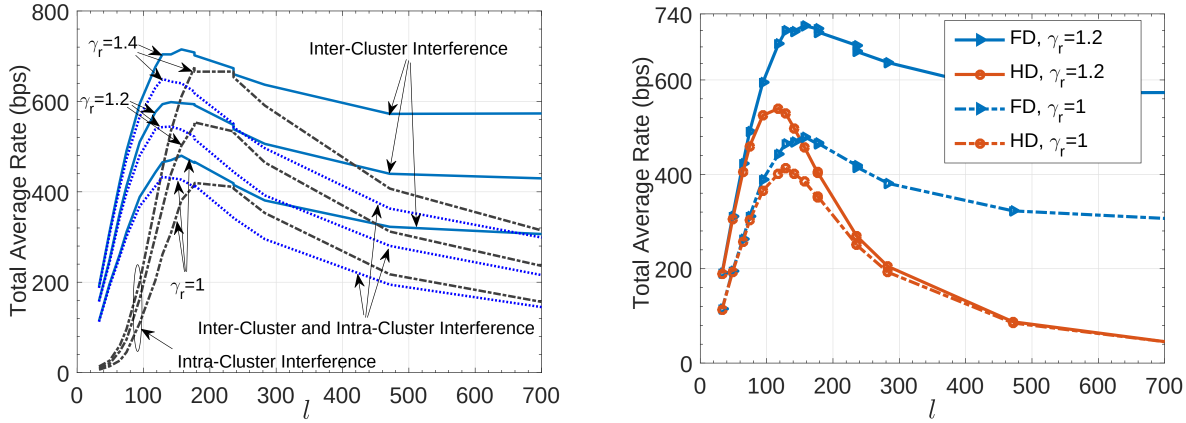
Fig. 4. *Expected number of links associated to HD/FD nodes, (a) Impact of Zipf exponent γ_r, (b) Impact of node density.*

optimisation problem in eq. (17) is depicted in Fig. 6(a). This figure illustrates the total average download time versus l. As can be seen from this figure, FD collaboration has major impact on decreasing the latency in downloading video files. The achievable reduction in delay is in the order of 45%.

Fig. 6(b) demonstrates the expected number of self-request users for different Zipf exponents. From Fig. 6(b), we observe that the number of self-request users increases as l decreases, and as the density of users within a cluster increases, the number of self-request users would decreases. This is because of the high probability of making a request from one user to the most popular files to make a D2D communication.

In the following simulations corresponding to Fig. 7, we have implemented the scenario of subsection 3.3 and obtained the numerical solutions of eq. (25). The values of the problem constraints are given in Tabel 1. Fig. 7 demonstrate tradeoff between the energy efficiency and the spectral efficiency in both HD-D2D and FD-D2D systems. As can be seen from this figure, FD-D2D mode outperforms HD-D2D mode in terms of spectral efficiency, while HD-D2D mode outperforms FD-D2D mode in terms of energy efficiency. The interpretations of such a tradeoff can be described as follows.

by increasing the cluster size, both parameters, i.e., $\Upsilon^{\delta}_{SE}(p^*_{ti}, p^*_{tj})$ and $\Upsilon^{\delta}_{EE}(p^*_{ti}, p^*_{tj})$, are decreasing, because, in lower cluster size, experienced SINR at respective receivers are of high value and consequently the spectral efficiency has its maximum value with optimal transmission powers, i.e., p^*_{ti} and p^*_{tj}. While in higher cluster size, the average of the experienced SINR at the respective receivers are low and consequently the spectral efficiency has its minimum value. In other hand, in lower cluster size, the optimal transmission powers due to short distances between D2D pairs, are of low value and the spectral efficiency has its maximum value, hence, the energy efficiency has its maximum value. while in higher cluster size, the optimal transmission powers due to longer distance between D2D pairs, are of high value and the spectral efficiency has its minimum value, hence, the energy efficiency has its minimum value. In the other words, in lower cluster size, the required transmission powers from the D2D transmitters to satisfy SINR thresholds at the respective receivers is low, while in higher cluster size, D2D transmitters need to use higher transmission powers to satisfy SINR thresholds at the respective receivers. Fig. 7 demonstrate that

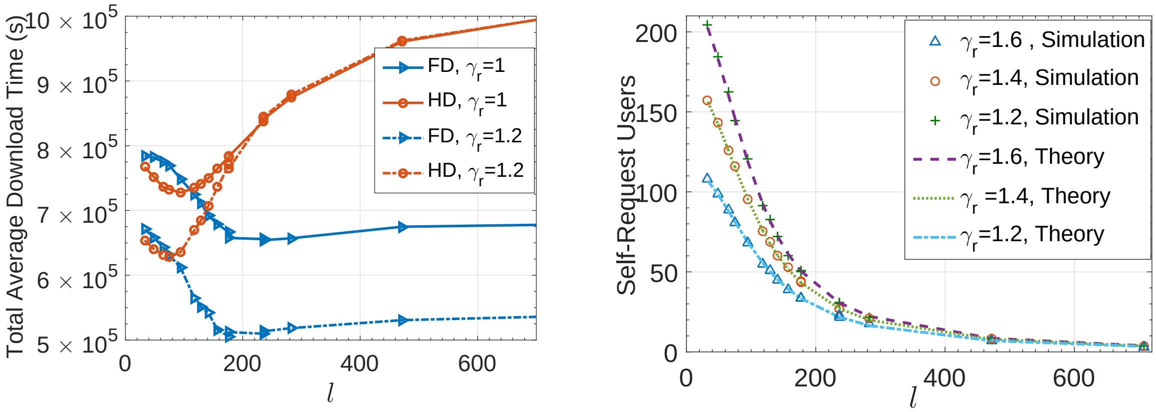


(a)

(b)

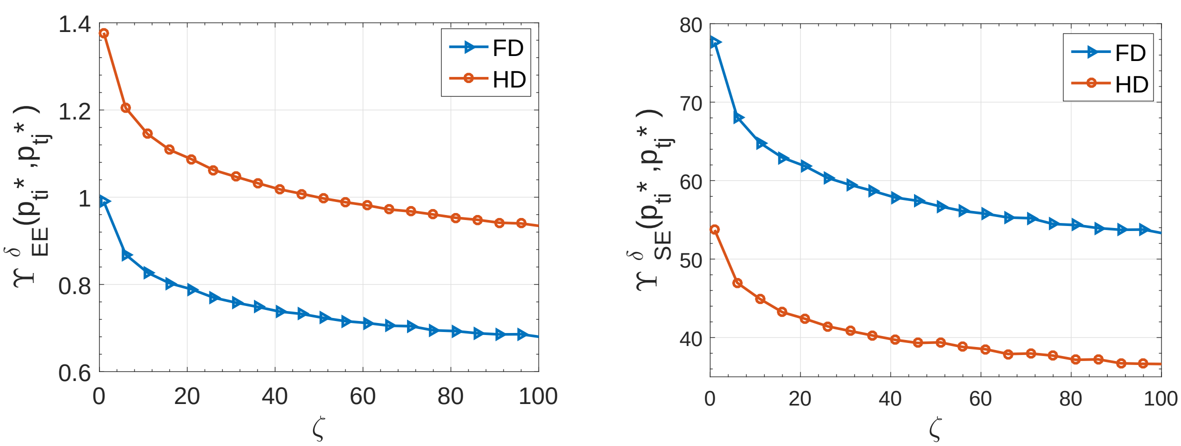
Fig. 5. *Total average rate versus D2D collaboration threshold* l*, (a) FD-only D2D network, (b) Comparison of FD-only and HD-only D2D network.*

FD video content delivery will require 40% higher energy consumption than HD, but will provide 44% highet SE. The achievable reduction in delay, as explained in Fig. 6(a) , will be in the order of 45%.



(a) (b)

Fig. 6. *Total average download time and self-request users versus l. (a) Total average download time. (b) Zero-latency users versus l.*



(a) (b)

Fig. 7. *Energy efficiency and spectral efficiency versus cluster size ζ, (a) Energy efficiency versus ζ, (b) Spectral efficiency versus ζ.*

5. CONCLUSIONS

In this paper, we used full duplex radios on user devices to increase the throughput and reduce delay of video caching in cellular systems with D2D collaboration. For the content placement, we used a priori content caching contents policy. We investigated a FD-enabled network by enabling FD radios only for D2D communications. Simulation results show that the achievable throughput can increase in high intra-cluster interference conditions. We also showed that allowing full duplex collaboration can have a major effect on the quality of video file distribution by reducing download time compared to HD-only collaboration. We also investigated the trade off between spectral efficiency and energy efficiency in FD and HD scenarios. We showed in a typical example that FD video content delivery will require 40% higher energy consumption than HD, but will provide 44% highet SE and 45% reduction in delay.

6. References

[1] 'Cisco',http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html, accessed in 11 December 2016

[2] Golrezaei, N., Molisch, A. F., Dimakis, A. G., Caire, G.: 'Femtocaching and device-to-device collaboration: A new architecture for wireless video distribution', IEEE Communications Magazine, 2013, vol. 51, pp. 142-149

[3] Shanmugam, K., Golrezaei, N., Dimakis, A. G., Molisch, A. F., Caire, G.: 'FemtoCaching: Wireless Content Delivery Through Distributed Caching Helpers', IEEE Transactions on Information Theory, 2013, vol. 59, pp. 8402-8413

[4] Ji, M., Caire, G., Molisch, A. F.: 'Fundamental Limits of Caching in Wireless D2D Networks', IEEE Transactions on Information Theory, 2016, vol. 62, pp. 849-869

[5] Golrezaei, N., Mansourifard, P., Molisch, A. F., Dimakis, A. G.: 'Base-Station Assisted Device-to-Device Communications for High-Throughput Wireless Video Networks', IEEE Transactions on Wireless Communications, 2014, vol. 13, pp. 3665-3676

[6] Bastug, E., Bennis, M., Debbah, M.: 'Living on the edge: The role of proactive caching in 5G wireless networks', IEEE Communications Magazine, 2014, vol. 52, pp. 82-89

[7] Afshang, M., Dhillon, H. S., Chong, P. H. J.: 'Fundamentals of Cluster-Centric Content Placement in Cache-Enabled Device-to-Device Networks', IEEE Transactions on Communications, 2016, vol. 64, no. 6, pp. 2511-2526

[8] Wu, Y., Chen, J., Qian, L. P., Huang, J., Shen, X.: 'Energy-Aware Cooperative Traffic Offloading via Device-to-Device Cooperations: An Analytical Approach', IEEE Transactions on Mobile Computing, 2016, vol. PP, pp. 1-1

[9] Yanru, Z., Lingyang, S., Saad, W., Dawy, Z., Zhu, H.: 'Contract-Based Incentive Mechanisms for Device-to-Device Communications in Cellular Networks', IEEE Journal on Selected Areas in Communications, 2015, vol. 33, pp. 2144-2155

[10] Lan, R., Wang, W., Huang, A., Shan, H.:'Device-to-Device Offloading with Proactive Caching in Mobile Cellular Networks', IEEE Global Communications Conference (GLOBECOM), 2015, pp. 1-6

[11] Saki, H., Shikh-Bahaei, M.: 'Cross-Layer Resource Allocation for Video Streaming Over OFDMA Cognitive Radio Networks', IEEE Transactions on Multimedia, 2015, vol. 17, pp. 333-345

[12] Knox, M. E.: 'Single antenna full duplex communications using a common carrier', Wireless and Microwave Technology Conference (WAMICON), 2012, 2012, pp. 1-6

[13] Xiufeng, X., Xinyu, Z.: 'Does full-duplex double the capacity of wireless networks?', IEEE Proc. INFOCOM, 2014, pp. 253-261

[14] Cheng, W., Zhang, X., Zhang, H.: 'Heterogeneous Statistical QoS Provisioning for Full-Duplex D2D Communications over 5G Wireless Networks', IEEE Global Communications Conference (GLOBECOM), 2015, pp. 1-7

[15] Dai, Z., Liu, J., Wang, C.: 'QoS-based device-to-device communication schemes in heterogeneous wireless networks', IET Communications, 2015, vol. 9, pp. 335-341

[16] Bobarshad, H., v. d. Schaar, M., Aghvami, H., Dilmaghani, R. S., Shikh-Bahaei, M. R.: 'Analytical Modeling for Delay-Sensitive Video Over WLAN', IEEE Transactions on Multimedia, 2012, vol. 14, pp. 401-414

[17] Xia, X., Xu, Y., Xu, K., Ma, W., Zhang, D.: 'Practical opportunistic full-/half-duplex relaying', IET Communications, 2015, vol. 9, pp. 745-753

[18] Cha, M., Kwak, H., Rodriguez, P., Ahn, Y. Y., Moon, S.: 'Analyzing the Video Popularity Characteristics of Large-Scale User Generated Content Systems', IEEE/ACM Transactions on Networking, 2009, vol. 17, pp. 1357-1370

[19] Yin, R., Zhong, C., Yu, G., Zhang, Z., *et al*.: 'Joint Spectrum and Power Allocation for D2D Communications Underlaying Cellular Networks', IEEE Transactions on Vehicular Technology, 2016, vol. 65, no. 4, pp. 2182-2195

[20] Chen, H., Li, G., Cai, J.: 'SpectralEnergy Efficiency Tradeoff in Full-Duplex Two-Way Relay Networks', IEEE Systems Journal , vol.PP, no.99, pp.1-10

[21] Chunguo, L., Cioffi, J. M., Yang, L.: 'Optimal energy efficient joint power allocation for twohop singleantenna relaying systems', Transactions on Emerging Telecommunications Technologies, 2014, pp. 745-751